

CLAIMS

1. (Previously Presented) A method of searching audio data, comprising:
receiving a query comprising a grammar corresponding to pronunciation alternatives that
define multiple phonetic possibilities for a segment of input speech; and
comparing the query with a lattice of phonetic hypotheses associated with the audio data
to identify if at least one of the multiple phonetic possibilities is approximated by
at least one phonetic hypothesis in the lattice of phonetic hypotheses.
2. (Previously Presented) The method of claim 1 and further comprising:
calculating a score indicative of the difference between said at least one of the multiple
phonetic possibilities and said at least one hypothesis in the lattice.
3. (Previously Presented) The method of claim 2 and further comprising:
adding a penalty value to the score if said at least one of the multiple phonetic
possibilities is approximated by inconsistent phonetic hypotheses.
4. (Previously Presented) The method of claim 1 and further comprising:
identifying a plurality of phonetic hypotheses in the lattice of phonetic hypotheses that
approximate the query.
5. (Previously Presented) The method of claim 4 and further comprising:
comparing the plurality of phonetic hypotheses identified.
6. (Previously Presented) The method of claim 1 and further comprising:
identifying a time span associated with said at least one phonetic hypothesis in the lattice
of phonetic hypotheses.

7. (Previously Presented) The method of claim 1 wherein the query is represented as at least one of a finite-state network, a context-free grammar and a prefix tree.
8. (Previously Presented) A method of generating a lattice from audio data, comprising:
recognizing phonetic fragments within the audio data wherein at least some of the phonetic fragments include at least two phones;
accessing a mutual information score for recognized phonetic fragments within the audio data that include at least two phones, wherein the mutual information score for each of the phonetic fragments having at least two phones is a function of a likelihood that phones in the phonetic fragment occur consecutively and a likelihood that each phone in the phonetic fragment occurs independent of other phones in the phonetic fragment; and
determining a score for paths joining adjacent phonetic fragments in the audio data using in part the mutual information score for the phonetic fragments having at least two phones.
9. (Previously Presented) The method of claim 8 and further comprising:
calculating time values and individual acoustic scores for each phone of each of the phonetic fragments.
10. (Previously Presented) The method of claim 8 and further comprising:
pruning paths that have a score that does not meet a threshold level.
11. (Previously Presented) The method of claim 8 and further comprising:
collapsing a plurality of audio frames in the audio data into a single audio frame.

12. (Previously Presented) A computer readable storage medium encoded with a data structure, comprising:

a plurality of phoneme hypotheses and an associated score for each hypothesis, wherein at least some of the hypotheses form phonetic fragments that include at least two phones, and wherein the score for each phonetic fragment that includes at least two phones is a function of a likelihood that phones in the phonetic fragment occur consecutively and a likelihood that each phone in the phonetic fragment occurs independent of other phones in the phonetic fragment; and
a plurality of transitions connecting the phoneme hypotheses.

13. (Previously Presented) The computer readable storage medium of claim 12 wherein the plurality of phoneme hypotheses corresponds to recorded audio data.

14. (Previously Presented) The computer readable storage medium of claim 13 wherein each associated score represents the likelihood of the phoneme hypotheses given the recorded audio data.

15. (Previously Presented) The computer readable storage medium of claim 12 wherein the plurality of transitions include an associated time value.

16. (Previously Presented) The computer readable storage medium of claim 12 wherein the data structure further includes a plurality of fragment transitions connecting one phoneme hypothesis in one phonetic fragment to another phoneme hypothesis in another phonetic fragment.